

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Geographical variations and factors associated with recent HIV testing prevalence in Ghana: spatial mapping and complex survey analyses of the 2014 demographic and health surveys
AUTHORS	Nutor, Jerry John; Duah, Henry; Duodu, Precious; Agbadi, Pascal; Alhassan, Robert; Darkwah, Ernest

VERSION 1 – REVIEW

REVIEWER	Larmarange, Joseph Centre Population et Développement, Institut de Recherche pour le Développement, Université Paris Descartes, ERL Inserm U 1244
REVIEW RETURNED	14-Dec-2020

GENERAL COMMENTS	<p>This article uses data from 2014 Ghana Demographic and Health Surveys to analyse factors associated with lifetime HIV testing (multivariate analysis using a Poisson model) and spatial variations (using kernel density estimators). Overall the paper is written correctly, the methods and results clearly presented. This is a good secondary analysis of such survey.</p> <p>I have, however, some methodological concerns.</p> <ol style="list-style-type: none">1. It seems that sampling weights have not been taken into account for the descriptive analysis (table 1) but also for the computation of the proportion of individuals who have ever tested for HIV in the manuscript.2. 50+ individuals have been included in the analysis. However, only the men sample included 50-59 years old, the women sample being limited to 15-49, according to the 2014 Ghana DHS report (https://dhsprogram.com/pubs/pdf/FR307/FR307.pdf). For comparability, it would have been better to limit the analysis to 15-49 only.3. The men sample and the women sample have been merged. It is problematic as only half of the eligible households were selected for the men survey while all households were selected for the women survey. Merging the two datasets would require to compute specific sampling weights (not available in the datasets) to account for these differences and to obtain correct estimates of the different indicators. The other option is not to merge the two datasets and to stratify all analysis. Note that this is the approach used in the official report. In chapter 13 presenting HIV results, no table combines men and women; all analyses are separate. For this paper, it will also simplify the presentation of the results. Currently, a first analysis combining men and women is presented before a stratified analysis. Presenting only stratified analysis will simplify the presentation of multivariate results.
-------------------------	---

	<p>4. The spatial analysis, in the absence of appropriate merging weights, is currently incorrect and should therefore be stratified by gender as well (and sampling weights being taken into account in prevR).</p> <p>5. For descriptive analysis (table 1), it would be more appropriate to present row percentage rather than column percentage, i.e. in each subgroup the proportion who ever tested for HIV.</p> <p>6. For multivariate analysis, considering the number of variables included, the authors could consider simplifying their models using a step approach (e.g. by minimising AIC). It could also be relevant to add global p-values to indicate which variables have a significant effect (ANOVA).</p> <p>7. I do not understand what are the pie charts in figure 1. Is it the proportion of men and women? In any case, it cannot be the prevalence by gender as pie charts are not appropriate for representing two proportions. That issue could be solved by stratifying the analysis by gender.</p> <p>Another point of discussion is the choice of analysing lifetime HIV testing. The authors could have chosen to look at recent testing (last 12 months or last 24 months), which is a better indicator of access to HIV testing. Indeed, lifetime testing is already low in Ghana, and recent testing is even lower: 13% of women tested in the last 12 months and 6% of men according to the DHS report. Although I recognise the additional amount of work, comparing both indicators (lifetime testing and recent testing) using the same approach (factor analysis and spatial analysis) would improve this paper and the discussion significantly.</p> <p>Additional comments: An alternative and more visual way of presenting the multivariate results could be to use a forest plot and to provide the full table as supplementary material. For kernel estimator spatial smoothing, it could be relevant to provide an idea of the level of smoothing (given by the smoothing circles radii). It could be added directly on the map (cf. Figure 15 of https://journals.openedition.org/cybergeog/24606) or in a separate figure in the supplementary material.</p>
--	--

REVIEWER	Yona, Sri Universitas Indonesia
REVIEW RETURNED	26-Jan-2021

GENERAL COMMENTS	Overall, the manuscript is good. Suggestion In the result section, the table for multivariate analysis can be made in simple way.
-------------------------	---

VERSION 1 – AUTHOR RESPONSE

Comments from the Editor:

- Please complete a thorough proofread of the text and correct any spelling and grammar errors that you identify.
- Please revise the title to indicate the research question, setting, and study design. This is the preferred format for the journal.
- Please ensure that all acronyms are defined on first mention, including those in the abstract.
- Please include any relevant statistical results in the results section of the Abstract.

- Along with your revised manuscript, please include a copy of the STROBE checklist indicating the page/line numbers of your manuscript where the relevant information can be found (<https://strobe-statement.org/index.php?id=strobe-home>).

Formatting Amendments (where applicable):

Response: Thank you for the relevant comments. We have proofread the text and corrected spelling and grammar errors. Also, we have revised the title according to your request (page 1), and edited our abstract to reflect the new results. We have added a STROBE checklist, too, as a supplementary file. Thank you

Reviewer: 1

Dr. Joseph Larmarange, Centre Population et Développement, Institut de Recherche pour le Développement, Université Paris Descartes, ERL Inserm U 1244

Comments to the Author:

This article uses data from 2014 Ghana Demographic and Health Surveys to analyse factors associated with lifetime HIV testing (multivariate analysis using a Poisson model) and spatial variations (using kernel density estimators).

Overall the paper is written correctly, the methods and results clearly presented. This is a good secondary analysis of such survey.

Response: Thank you for the positive feedback. We are grateful.

I have, however, some methodological concerns.

1. It seems that sampling weights have not been taken into account for the descriptive analysis (table 1) but also for the computation of the proportion of individuals who have ever tested for HIV in the manuscript.

Response: Thank you for the comments. We used the `svyset` command and adjusted for stratification, primary sampling units, and sample weights for all analyses. See Table 1 on pages 12-13.

2. 50+ individuals have been included in the analysis. However, only the men sample included 50-59 years old, the women sample being limited to 15-49, according to the 2014 Ghana DHS report (<https://dhsprogram.com/pubs/pdf/FR307/FR307.pdf>). For comparability, it would have been better to limit the analysis to 15-49 only.

Response: We used the `subpop` function associated with the `svy` command to limit the analyses for the men to 15-49 years only, and this has been documented in the methods section on page 6.

3. The men sample and the women sample have been merged. It is problematic as only half of the eligible households were selected for the men survey while all households were selected for the women survey. Merging the two datasets would require to compute specific sampling weights (not available in the datasets) to account for these differences and to obtain correct estimates of the different indicators. The other option is not to merge the two datasets and to stratify all analysis. Note that this is the approach used in the official report. In chapter 13 presenting HIV results, no table combines men and women; all analyses are separate. For this paper, it will also simplify the presentation of the results. Currently, a first analysis combining men and women is presented before a stratified analysis. Presenting only stratified analysis will simplify the presentation of multivariate results.

Response: Thank you for the comment. We have simplified the analyses by presenting only stratified results from the multivariable model. This is reported in S1 Table; at the end of the reference list on page 27.

4. The spatial analysis, in the absence of appropriate merging weights, is currently incorrect and should therefore be stratified by gender as well (and sampling weights being taken into account in prevR).

Response: Thank you for the comment. We have produced separate surface maps for women and men, and we accounted for the sample weight in prevR.

5. For descriptive analysis (table 1), it would be more appropriate to present row percentage rather than column percentage, i.e. in each subgroup the proportion who ever tested for HIV.

Response: Thank you for the comment. We have reported row percentages for the cross-tab. Table 1 has this information on page 12-13.

6. For multivariate analysis, considering the number of variables included, the authors could consider simplifying their models using a step approach (e.g. by minimising AIC). It could also be relevant to add global p-values to indicate which variables have a significant effect (ANOVA).

Response: Thank you for the comments. First, we generated global p-values for all the variables under consideration in our study. Secondly, The “gvselect”—Best subsets variable selection—in STATA was used to identify the best features to build models that explain the variability in HIV testing among Ghanaian men and women samples. Details of these are reported on pages 8-9.

7. I do not understand what are the pie charts in figure 1. Is it the proportion of men and women? In any case, it cannot be the prevalence by gender as pie charts are not appropriate for representing two proportions. That issue could be solved by stratifying the analysis by gender.

Response: Thank you for this comment. As earlier indicated, we stratified the analyses by gender.

Another point of discussion is the choice of analysing lifetime HIV testing. The authors could have chosen to look at recent testing (last 12 months or last 24 months), which is a better indicator of access to HIV testing. Indeed, lifetime testing is already low in Ghana, and recent testing is even lower: 13% of women tested in the last 12 months and 6% of men according to the DHS report. Although I recognise the additional amount of work, comparing both indicators (lifetime testing and recent testing) using the same approach (factor analysis and spatial analysis) would improve this paper and the discussion significantly.

Response: Thank you very much for the comment. We limited our analyses to recent HIV testing (last 12 months).

Additional comments:

An alternative and more visual way of presenting the multivariate results could be to use a forest plot and to provide the full table as supplementary material.

For kernel estimator spatial smoothing, it could be relevant to provide an idea of the level of smoothing (given by the smoothing circles radii). It could be added directly on the map (cf. Figure 15 of <https://journals.openedition.org/cybergeogeo/24606>) or in a separate figure in the supplementary material.

Response: Thank you for the comments. We have, as suggested, produced the surface maps with the smoothing circle radius contours (in kilometers) and have added as supplementary figures (Women: S1 Figure, Men: S2 Figure).

Reviewer: 2

Dr. Sri Yona, Universitas Indonesia

Comments to the Author:

Overall, the manuscript is good. Suggestion In the result section, the table for multivariate analysis can be made in simple way.

Response: Thank you for the comments. We have addressed these comments and have simplified our multivariable tables. Thank you.

VERSION 2 – REVIEW

REVIEWER	Larmarange, Joseph Centre Population et Développement, Institut de Recherche pour le Développement, Université Paris Descartes, ERL Inserm U 1244
REVIEW RETURNED	03-May-2021

GENERAL COMMENTS	<p>This new version of the manuscript has been significantly improved. The analysis is now focused on recent HIV testing (less than 12 months) and completely stratified by gender. The authors answered to all comments sent to them.</p> <p>I just have minor suggestions for presentation improvements/clarifications:</p> <ul style="list-style-type: none"> • It could be worth adding in the title 'recent' before 'HIV testing', i.e. "factors associated with recent HIV testing" • Table 1: the fact that results are survey-weighted should be mentioned in the title or in a table note. • Table 1: some elements are unclear and the table could more explicit and simplified: (i) replace 'total' by 'number of observations' (I assume that this is the real number of observations and not the weighted numbers); (ii) present the distribution in a separate column 'weighted percentage of the total population'; (iii) remove the 'No' column; (iv) replace 'yes' by 'HIV test in the past 12 months (weighted %)' • What are the p-values in this table? Add a table note. Should not be in the 'total' column (not explicit) but rather in a dedicated column. • Figure 4 and 5: increase legend size. It is difficult to read. • Tables S1: (i) define APR in a table note; (ii) what is the difference between the two columns. I understand that the second one presents the reduced model. But what is the first column? The full model? Or the bivariate analysis? Please explain the last column (adjusted Wald test)? What is tested?
-------------------------	--

	<ul style="list-style-type: none"> • Supplementary figures: increase the text size. It is very difficult to read.
--	--

VERSION 2 – AUTHOR RESPONSE

Response to reviewers

Comments: This new version of the manuscript has been significantly improved. The analysis is now focused on recent HIV testing (less than 12 months) and completely stratified by gender. The authors answered to all comments sent to them.

Response: Thank you very much for your earlier and current comments and suggestions.

I just have minor suggestions for presentation improvements/clarifications:

Comment: It could be worth adding in the title 'recent' before 'HIV testing', i.e. "factors associated with recent HIV testing"

Response: We have revised the title in line with your comment. Thank you.

Comments on Table 1:

- Table 1: the fact that results are survey-weighted should be mentioned in the title or in a table note.
- Table 1: some elements are unclear and the table could more explicit and simplified: (i) replace 'total' by 'number of observations' (I assume that this the real number of observations and not the weighted numbers); (ii) present the distribution in a separate column 'weighted percentage of the total population'; (iii) remove the 'No' column; (iv) replace 'yes' by 'HIV test in the past 12 months (weighted %)'
- What are the p-values in this table? Add a table note. Should not be in the 'total' column (not explicit) but rather in a dedicated column.

Response: Thank you for these comments on Table 1. The revised Table 1 reflects all the required changes.

Comment: Figure 4 and 5: increase legend size. It is difficult to read.

Response: Thank you. The newly plotted figures 4 and 5 have large and easy to read legends.

Comment: Tables S1: (i) define APR in a table note; (ii) what is the difference between the two columns. I understand that the second one presents the reduced model. But what is the first column? The full model? Or the bivariate analysis? Please explain the last column (adjusted Wald test)? What is tested?

Response: Thank you for these comments on Table S1. The revised Table S1 reflects all the required changes. The last column has been deleted since that section has already been explained in the main manuscript on pages 8 and 14.

Comment: Supplementary figures: increase the text size. It is very difficult to read.

Thank you. The newly plotted supplementary figures have large and easy to read texts.